DOCUMENT RESUME

ED 028 682 48 FL 001 325

By-Keiler, Allan

Surface Structure and Deep Structure in Latin Syntax. Studies in Language and Language Behavior, Progress Report VIII.

Michigan Univ., Ann Arbor. Center for Research on Language and Language Behavior.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Report No-SLLB-PR-8
Bureau No-BR-6-1784
Pub Date 1 Feb 69
Contract-OEC-3-6-061784-0508
Note-15p.
EDRS Price MF-$0.25 HC-$0.85
Descriptors-*Deep Structure, English, Instructional Innovation, Language Research, *Latin, Linguistics, Linguistic Theory, Phrase Structure, Sentence Diagraming, Sentence Structure, Structural Analysis, Structural Grammar, *Surface Structure, *Syntax, Teaching Methods, Transformation Generative Grammar, Transformations (Language), *Transformation Theory (Language), Verbs

In an attempt to apply recent developments in transformational grammar to Latin syntax, this report analyzes first English, then Latin sentences for both deep and surface structures through transformational and phrase structure grammar methods. Auxiliary nodes, problems of Latin verb complimentation, and the gerund and gerundive constructions are the syntactical problems in Latin considered. (AF)

ED028682

BR-6-1784
PA-48
OE-BR

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SURFACE STRUCTURE AND DEEP STRUCTURE IN LATIN SYNTAX[1]

Allan Keiler

Center for Research on Language and Language Behavior
The University of Michigan

The current notion of deep structure versus surface structure is examined from the point of view of Latin syntax. The syntactical problems in Latin dealt with are the auxiliary node, some problems of Latin verb complimentation, and the gerund and gerundive constructions.

Within the last decade an increasing amount of work on the description of natural languages has been carried out within the framework of generative or transformational grammar. Latin, however, has been the subject of only a few studies in terms of transformational grammar,[2] at least in comparison with studies of other languages. And none of the studies has yet taken account of any of the significant changes that transformational grammar has undergone in the past several years.[3] It may be useful to characterize some of the contributions of these recent developments in grammatical description by showing how they might apply to the analysis of Latin syntax.

It will perhaps be most profitable to concentrate on a single essential characteristic of recent generative grammar--the distinction between deep structure and surface structure. Taxonomic linguistics, i.e., the kind of linguistic description prevalent in the United States from Bloomfield on, limited the kinds of statements made about language structure to the overt form of sentences, or to the grammatical elements that actually occur in a sentence; to the order and morphological shape of those elements; and to their actual phonological form. Syntactic descriptions of this kind have been variously referred to as immediate constituent analysis or phrase structure grammar, and are more or less similar to traditional sentence parsing. Thus the sentence

(1) The detective shot the bandit.

would have the following kind of syntactic description, in terms of the hierarchy of its constituent-construction relationships:

FL 001 325

FL001325

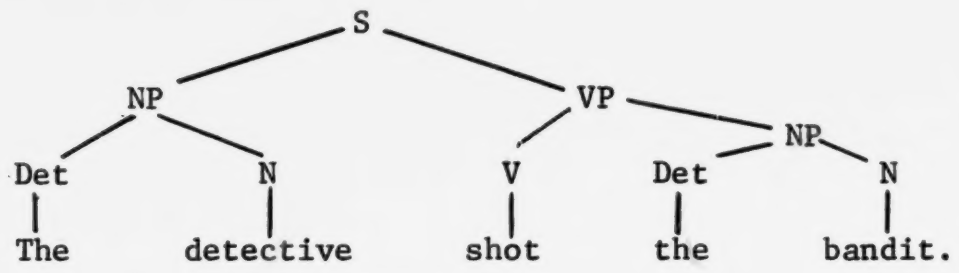


Figure 1

A diagram such as this is called a phrase marker or immediate constituent tree diagram, and results automatically from a set of syntactical rules with limited and well-defined properties. These rules are in the form of simple rewriting rules, in which single constructions are rewritten as the sequence of their immediate constituents until the ultimate constituents of the sentence are reached. Figure 1 is derived, therefore, from such rules as

(i) S → NP + VP

(ii) NP → Det + N

etc., where S is Sentence, NP is Noun Phrase, VP is Verb Phrase, Det is Determiner, and N is Noun. Thus both the paradigmatic or part-of-speech features, and syntagmatic, or co-occurrence relationships, of the sentence are formally represented by such a diagram. In terms of co-occurrence relationships, for example, the phrase marker indicates the degree of grammatical intimacy which obtains between any two or more constituents of a sentence by allowing them to be traced back to their lowest dominating node: detective and shot are thus ultimate constituents of the sentence only, whereas the and bandit form a NP, which in turn forms, together with V, a VP. Such phrase markers are insightful only to the degree that all correct syntactical relationships are manifested by the overt form of sentences.[4]

The analysis of the verb phrase in both English and Latin will serve as an instructive initial example of the insight gained in distinguishing the deep structure from the superficial structure of sentences.[5] In English the active verb phrase is exemplified by the following: (he) eats, is eating, has eaten, has been eating, can eat, can have eaten, can have been eating, etc., and in the past, (he) ate, was eating, had eaten, etc. Or in general terms, the verb phrase must obligatorily include Tense and MV (Main Verb) and may optionally include a Modal, as well as signals for the progressive or perfective. Now the generalities underlying these examples as a whole are obscured by two obligatory features of the overt form of these verb phrases: the "movable" suffixation of participles and tense, and the discontinuity of the auxiliaries be and have from the suffixes for the past and

present participles. On the one hand, tense may be signaled by any of the verbal stems: Modal (could have eaten), have (had eaten), be (was eating) or MV (ate); and on the other hand, have and past participle suffix -en, and be and present participle suffix -ing, although constituents of a single construction, are never contiguous in the verb phrase. Abstracting these two overt features of verb phrases, Figure 2 provides the correct constituent relationships for English:

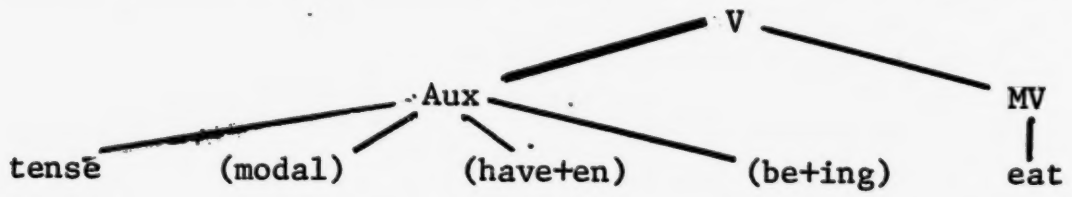


Figure 2

In other words, every type of verb phrase must conform to the string Tense (Modal) (have+en) (be+ing) MV, in that order, where Modal, have+en and be+ing are optional elements and Tense and MV, obligatory elements, and these terminal elements have the constituent structure indicated in Fig. 2. Further, the co-occurrence relationships between have and past participle and be and present participle are made explicit since they are now formalized as immediate constituents of the same construction.

Notice that the phrase marker in Fig. 2 provides all the syntactical information necessary to describe correctly the verb phrases in question, but it is not identical with the examples given above. To derive these examples, i.e. the overt form of verb phrases or the basis for their pronunciation, another level of analysis is required, with different properties from those of the phrase structure. This additional level is the transformational level, which may rearrange or delete elements according to prescribed limitations.[6] The overt form of the verb phrase in English must be derived automatically by a transformational rule which suffixes any member of the affix class (tense, en, ing) to the next member of the stem class (Modal, have, be, MV). To take a single example, could have been eating is derived from its underlying phrase marker in the following way

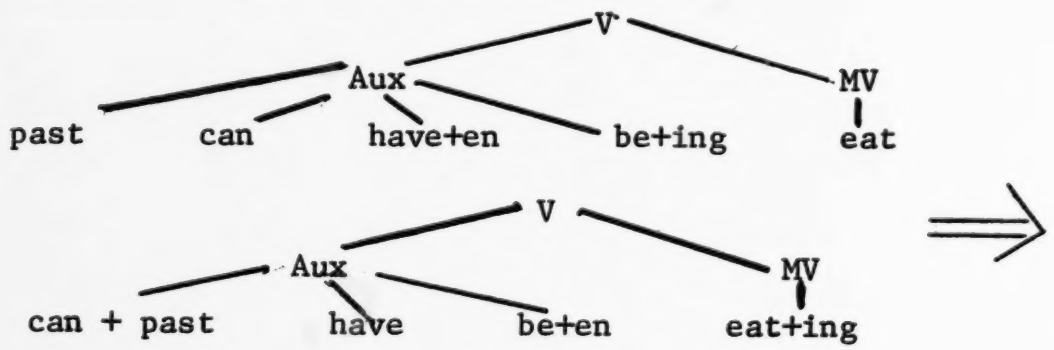


Figure 3

We thus distinguish between deep or underlying phrase markers, which make explicit all of the correct syntactical relationships of sentences, and which are the basis for their proper semantic interpretation, from superficial or derived phrase markers, which are the basis for the correct phonological and inflectional implementation of sentences.

Now the Latin verb phrase is also characterized by the same type of "moveable" suffixation (quite apart from the particular order of words ultimately chosen in the verb phrase) as well as by discontinuity. Consider first venit, potest venire. Either MV or Modal may signal tense (we may here omit considerations of person, number, etc.) and if Modal is chosen, the infinitive form of the verb must appear, i..e. there is mutual dependence between Modal and re (used here to represent the appropriate infinitival form) in the same way as between, e.g., have + en in English. The abstract generality underlying all such examples in Latin is accordingly that of Verb Tense (Modal + re) MV, where both Tense and re become automatically suffixed to the next chosen element (Modal or MV) by the same type of automatic switching rule as in English.

The underlying or deep structure of the Latin verb phrase which includes perfective aspect is easier seen if the main verb or modal is deponent rather than non-deponent. Consider such examples as, secùtus erat (with MV only) or cōnātus erat venire (with both Modal and MV). In both cases the perfective aspects necessarily includes two constituents which are not contiguous, i.e. with one suffixed to the other: some form of the verb esse (here represented by es) and some form of the perfect passive participle, and these constituents are mutually dependent. In the first example, however, the MV (sequor) is in the form of the perfective passive participle and in the second, the Modal (cōnor) is in the participial form. Both types of examples can be derived from an underlying verb phrase with the following deep structure: Tense (es + tus) (Mod + re)MV. Secūtus erat is derived from Past (es + tus) sequ- with both past and tus suffixed to the following stem; cōnātus erat venire is, accordingly, derived from past (es + tus) (cōnā + re) veni-, with past, tus, and re suffixed to the following stem.

Consider now examples such as potuerat venire. vēnerat, etc., i.e. verb phrases with non-deponent Modals or MV's. Now secūtus erat and vēnerat, on the one hand, as well as potuerat venire and cōnātus erat venire, on the other, clearly have the same underlying syntactical description. In other words, in spite of different morphological components, both pairs of examples must be

described as containing the same grammatical formatives. This is similar to claiming that, e.g., worked and took in English are different inflectional implementations of the same underlying grammatical formatives, past + verb. Hence, the first pair of Latin examples above is derived as well from past (es + tus) MV, the second pair from past (es + tus) (Mod + re) MV. The implementation of the relevant formatives for non-deponents in Latin is, as we shall see, just more removed from the underlying structure than is the case for deponents, i.e. will require an additional transformation rule.

The following phrase markers, then, underlie venerat and potuerat venire:

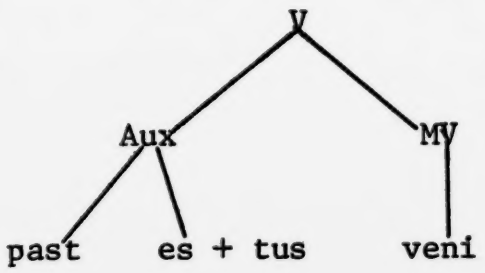


Figure 4

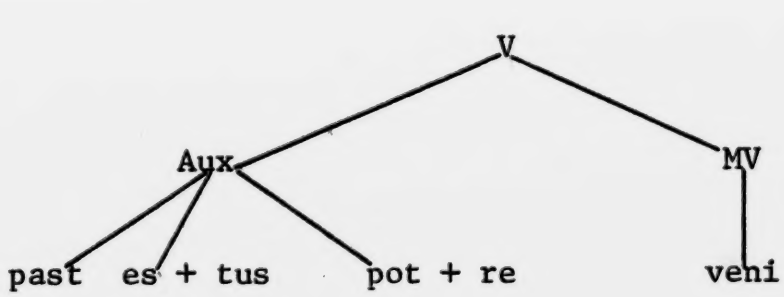


Figure 5

By applying the same suffix switching rule to the above examples, the following representations are derived from the above phrase markers.

i) (es + past) (veni + tus)

ii) (es + past) (pot + tus) (veni + re)

But since we are now dealing with non-deponent MV's and modals, it is sufficient to add an additional switching rule which attaches the (es + past) constituent to the following one, deriving:

i) (veni + tus + es + past)

ii) (pot + tus + es + past) (veni + re)

All of the inflectional structure of the verb phrases can now be derived automatically: when tus appears as the last formative after a stem, the stem is implemented as the perfective stem. Thus (pot + tus + es + past) is implemented

as potuerat, etc. This is merely to claim that tense and aspect are implemented for deponents as two free forms, but for non-deponents as one, which, in fact, is the case for Latin.

To summarize, then, all verb phrases in Latin are described as a selection from the following underlying string:

Tense (es + tus) (Mod + re) MV

The superficial form of verb phrases is then derived automatically by the affix switching rule which switches any occurrence of Affix (tense, tus, or re) with any occurrence of Stem (es, Mod, MV). For non-deponents, the first constituent i.e. (es + tus), is again switched to the following constituent, and the resulting sequences can then be implemented further by a set of phonological rules. In this way every example of a Latin verb phrase conforms to the same abstract underlying representation, but this insight is possible only if the overt form of the verb phrase is not made the sole basis of the description. Furthermore, it is important to note that the result of this claim is that the structural order of elements in the above string is a necessary part of the description, i.e., the four formatives must be described in the proper linear order; otherwise, the suffixes will become attached incorrectly to stems. In other words, it will turn out that "free" word order is a part of the superficial aspect of Latin syntax only, but not of the underlying form of sentences. In fact, research in generative grammar has made it increasingly clear that fixed order is universally a property of the underlying structure of language, although individual languages may differ in their superficial structure in terms of fixed vs. free word order.[7]

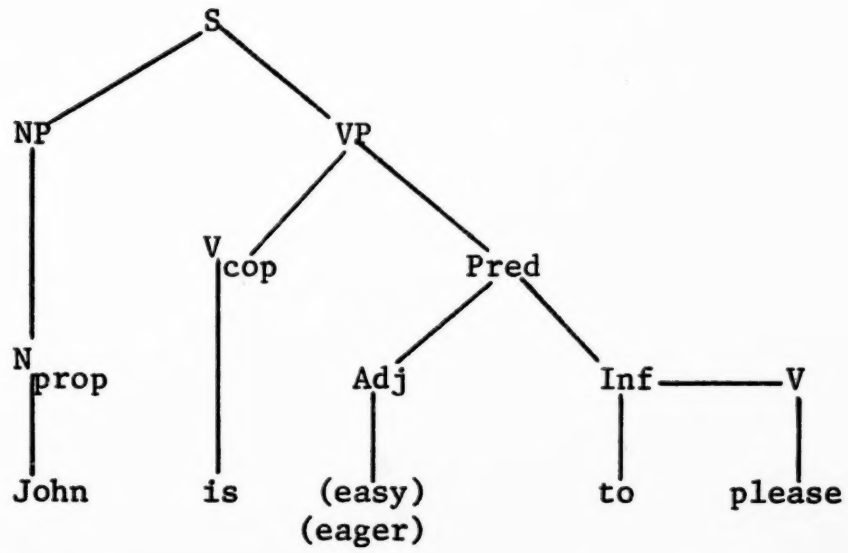
Consider, next, the following pair of English sentences:

2) John is eager to please.

3) John is easy to please.

The phrase marker for both of these, given the overt appearance of these sentences, will turn out to be exactly the same. That is, (2) and (3) can be analyzed into immediate constituents in the same way only, and this information is represented in Figure 6. But clearly these sentences are syntactically different in at least one important way. In (2) John is understood as both the subject of is (eager) as well as please, whereas in (3) John is understood simultaneously as the subject of is easy (at least the "grammatical" subject), and also as the complement of please. Thus a grammar which does not

Figure 6



distinguish (2) and (3) is incorrect. If we consider, then, the above phrase marker to be a more or less correct syntactical representation of (2), it will be necessary to provide a different representation for (3) which will, among other things, make explicit the differences between (2) and (3).

Notice that the only distinction between (2) and (3) is between the adjectives easy and eager, but that it is not possible to distinguish the phrase markers for (2) and (3) on the basis of the adjectives (or adjective classes), since easy and eager have in these sentences the same syntactical environments. Now compare (3) with the following pair of sentences,

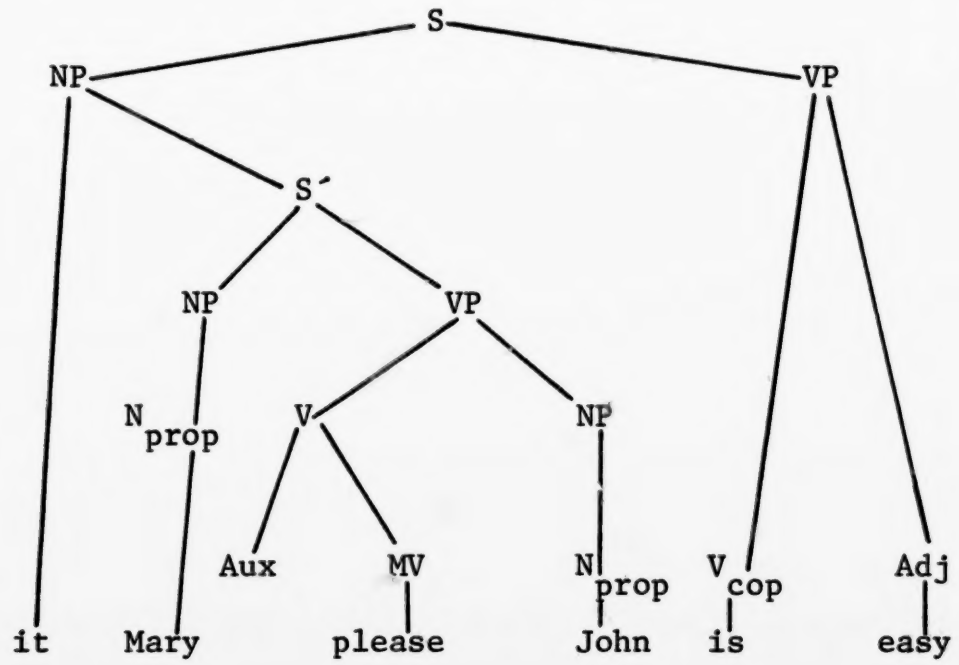
(3i) It is easy to please John.

(3ii) To please John is easy.

all three of which have the same semantic interpretation (i.e., mean essentially the same thing, matters of emphasis aside). Furthermore, the verb complement relationship between please and John is more easily describable for (3i) and (3ii) since the two constituents are contiguous in these sentences. Notice that it is possible additionally to have a subject of please John in these sentences: It is easy for Mary to please John, and For Mary to please John is easy. (3i) is different superficially from (3ii) to the extent that the real subject of is easy follows the predicate, and is replaced by the grammatical subject it. There are, therefore, two sentences which make up all of the examples in (3),

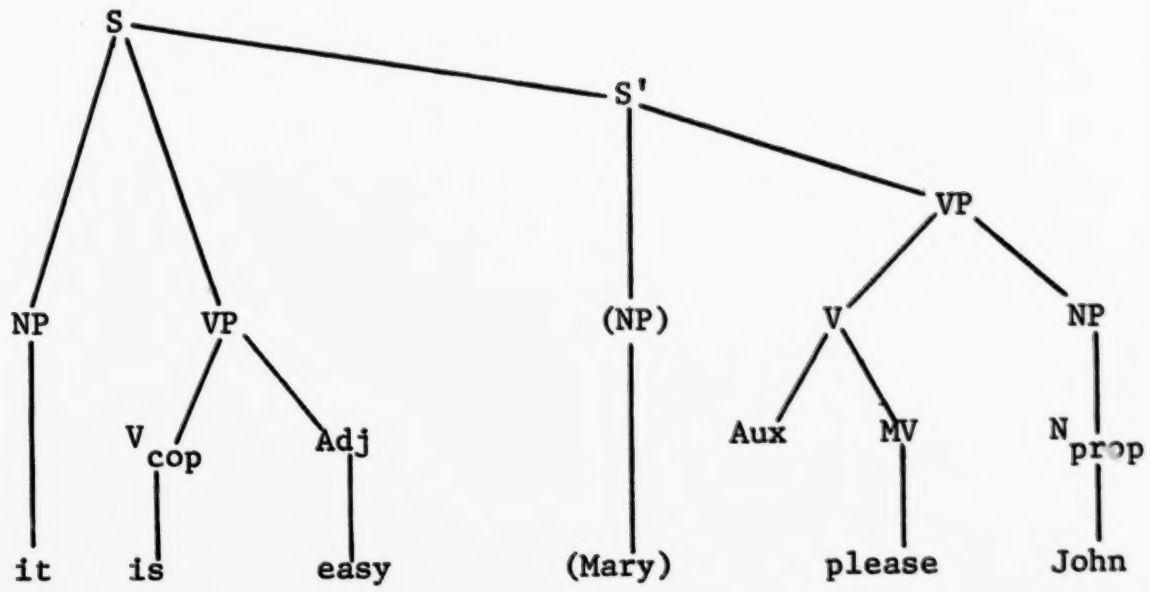
one of which (Mary pleases John) is embedded in the other ('something' is easy). The embedded sentence in these examples functions as the subject of the matrix sentence (x is easy). (3, 3i, 3ii) are all different syntactical implementations of the same underlying phrase marker. Now by abstracting these superficial differences, one can generalize about the underlying syntactical identity of (3), (3i) and (3ii) by means of the following phrase marker:

Figure 7



The subject of both (3i) and (3ii) is, then, described as a nominalization of a sentence whose overt form is determined by its position in the phrase marker in which it is embedded, but whose underlying syntactical relationships are identical to those of the non-embedded form, in this case, Mary + Aux + please + John. Both (3i) and (3ii) are then derivable from the above phrase marker by the appropriate transformational rules: (3ii) by obligatory deletion of it and optional deletion of the subject NP of the embedded sentence (S'), and (3i) by the rule which attaches the S' as a further constituent of S.[8] If the second alternative is chosen, then one has the following derived phrase marker: (3) is then derivable from the rule which substitutes John for it in Figure 8, or in general, which substitutes the verbal complement of the embedded sentence for it. Adjectives like easy can now be distinguished from eager: (is) easy may co-occur with embedded sentence subjects, but (is) eager may not, since *to please John is eager is, of course, ungrammatical. Thus the ambiguity of (2) and (3) is made the result of the application of transformational rules

Figure 8



which convert an underlying, fully explicit phrase marker into the derived phrase marker represented by the overt shape of these sentences.

To take a somewhat similar example from Latin, now, consider the following sentences:

(4) Caesar iussit mīlitēs castra pōnere.

(5) Caesar dīxit mīlitēs castra pōnere.

The immediate constituent analysis or phrase markers of these sentences is similarly ambiguous, i.e., both can be cut into immediate constituents in only one way, represented in Figure 9.

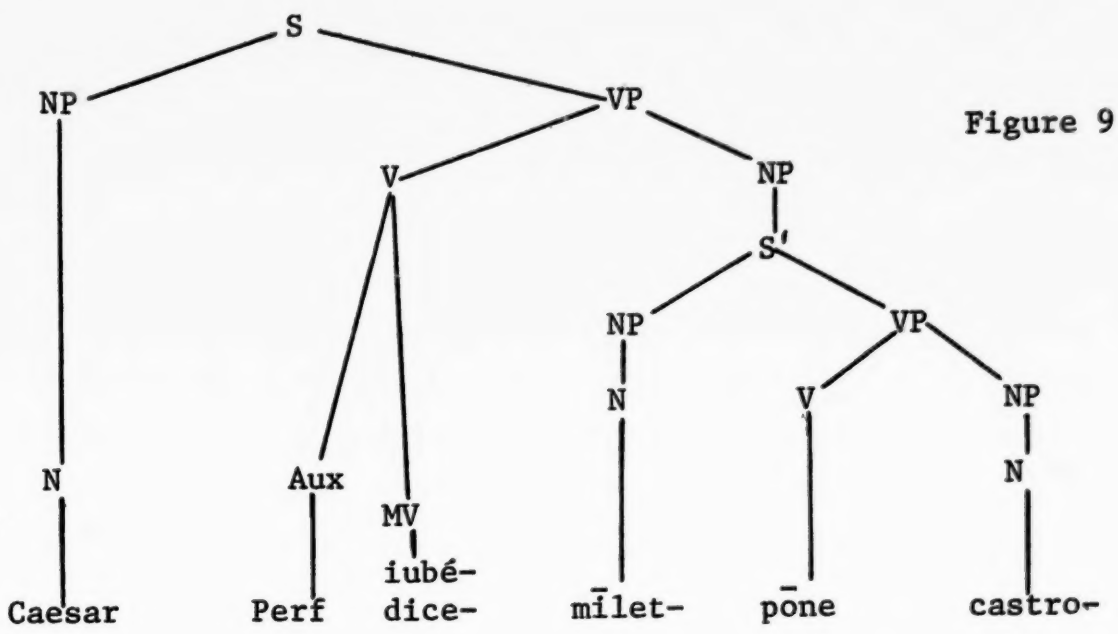


Figure 9

Now Figure 9 is correct for (5), since the embedded sentence must be considered a sentential complement, which becomes obvious in the passive sentence derived from (5), namely

(6) mīlitēs castra pōnere dictum est

where the S' has become the subject of dictum est, hence the impersonal passive construction.[9] If the same phrase marker is given for (4), however, then a passive construction identical to (6) would be derived, namely

(7) *mīlitēs castra pōnere iussum est

which is non-grammatical, the only possible passive being mīlitēs iussī sunt castra pōnere. The passives of these sentences, among other things, show clearly that the phrase marker for (4) cannot be the same as the one for (5), although the overt form of the latter permits no other phrase marker. Kühner,[10] e.g., perceptively argues that verbs of the class of iubeō are different from dicō, since "In Sätzen wie iubeō tē venīre, sinō tē redīre lässt sich der Akkusativ nicht nur als Object des regierenden Verbs...., sondern zugleich auch als Subject des abhängigen Infinitivs auffassen." The passive sentences associated with (5), in fact, make it clear that in the underlying phrase marker for (5), mīlitēs but not mīlitēs castra pōnere must be the verbal complement, hence the "personal" passive construction. We can then provide the following phrase marker for (5):

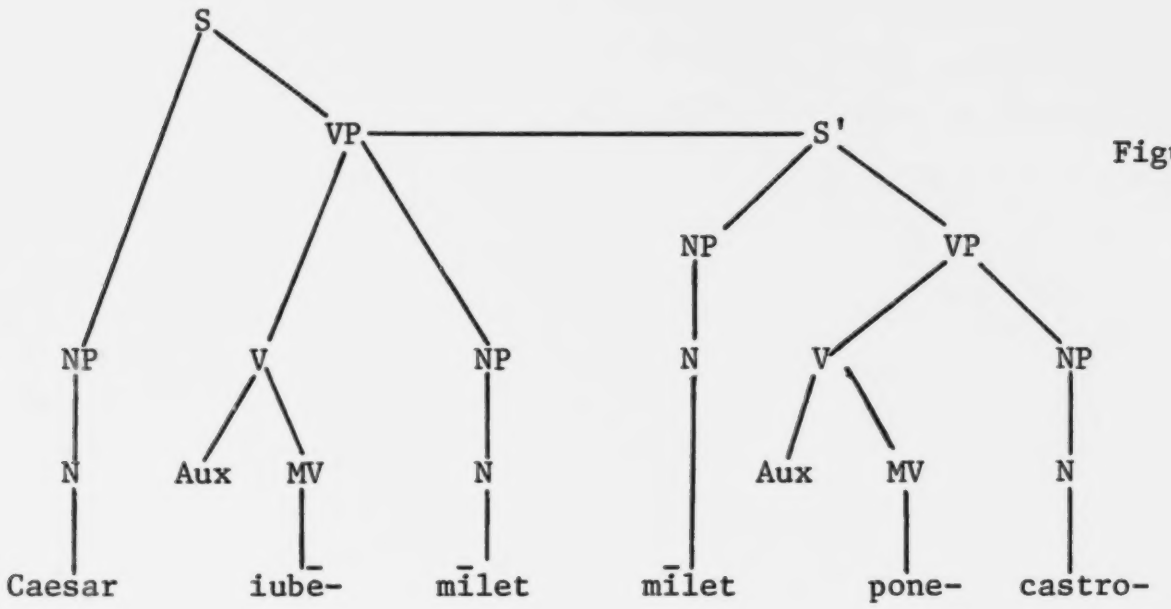


Figure 10

The above phrase marker generalizes about verbs of the iubeō class, then, by specifying the constituents of the V to be a personal N + S' (i.e. embedded

sentence), and further, that this personal N and the NP subject of the embedded sentence be identical, i.e., that there be co-reference between these NP's. The passive transformation will then be automatically personal (since the NP, but not the S', is the verbal complement), and the double function of the personal noun in the environment of such verbs noted by Kühner becomes explicit by insisting on co-reference. (5) is derived from the underlying phrase marker by deletion of the repeated NP and substitution of the Aux by re, which will become automatically attached to the following verb stem by the affix switching rule already discussed. The ambiguity that (5) shares with (4) on the level of immediate constituents thus becomes the automatic result of the application of general transformational rules which convert the more explicit underlying phrase marker to its overt or superficial shape.

To take a final example, let us consider briefly the problem of the gerund and gerundive in Latin. Halle and Buck[11] present what seems to me an entirely successful account of how the gerund, gerundive (and future passive participle) ought to be described, although they must consistently face the difficulties of limiting their descriptions to the inflectional form of the constructions in question, that is, to their superficial structure. They define gerund as a complete verbal noun, the gerundive as nearly the equivalent of a verbal noun. The latter definition is, of course, embarrassingly equivocal: either something is or is not a verbal noun. The point is, however, that apart from their different overt shapes, gerund and gerundive constructions have the same meaning. Furthermore, they are identical in the features which oppose them to the true future passive participle (e.g. in Carthāgō delēnda est but not in spēs Carthāginis dēlendae). Accordingly, Hale and Buck describe the gerund and gerundive, as opposed to the true future passive participle, as:

A) expressing the leading idea of their phrase
B) conveying no idea of necessity or obligation
C) active in feeling, not passive
D) accordingly unable to take any construction of the agent.

A asserts that gerundives have the same semantic interpretation as their equivalent gerund constructions, even though the gerundive is not the leading idea of the phrase-dē contemnendā glōriā, e.g., = dē contemptiōne glōriae both grammatically and semantically. The description of the gerundive as passive, then,

is made necessary because of the overt shape of such constructions in which the gerundive modifies a noun and must therefore be opposite in voice to gerunds (which govern the same noun) if their semantic interpretations are to come out as identical. (C) and (D) provide strong support, in fact, for suggesting that the description of gerund and gerundive as active and passive is misleading, if not incorrect.

All of the difficulties in Halle and Buck's treatment of the gerund and gerundive disappear if the underlying phrase marker for both gerund and gerundive constructions is considered a nominalization of a necessarily active sentence. Here again, by nominalization, is meant the occurrence of a sentence embedded within another sentence and functioning in that sentence (usually termed matrix sentence) in place of a noun, and whose overt shape is determined by its position in the matrix phrase marker. The phrase marker underlying both

(8) Poetae in iīs libellis dē contemnendā glōriā scrībunt

(9) Poetae in iīs libellīs dē contemnendō glōriam scrībunt

is, accordingly:

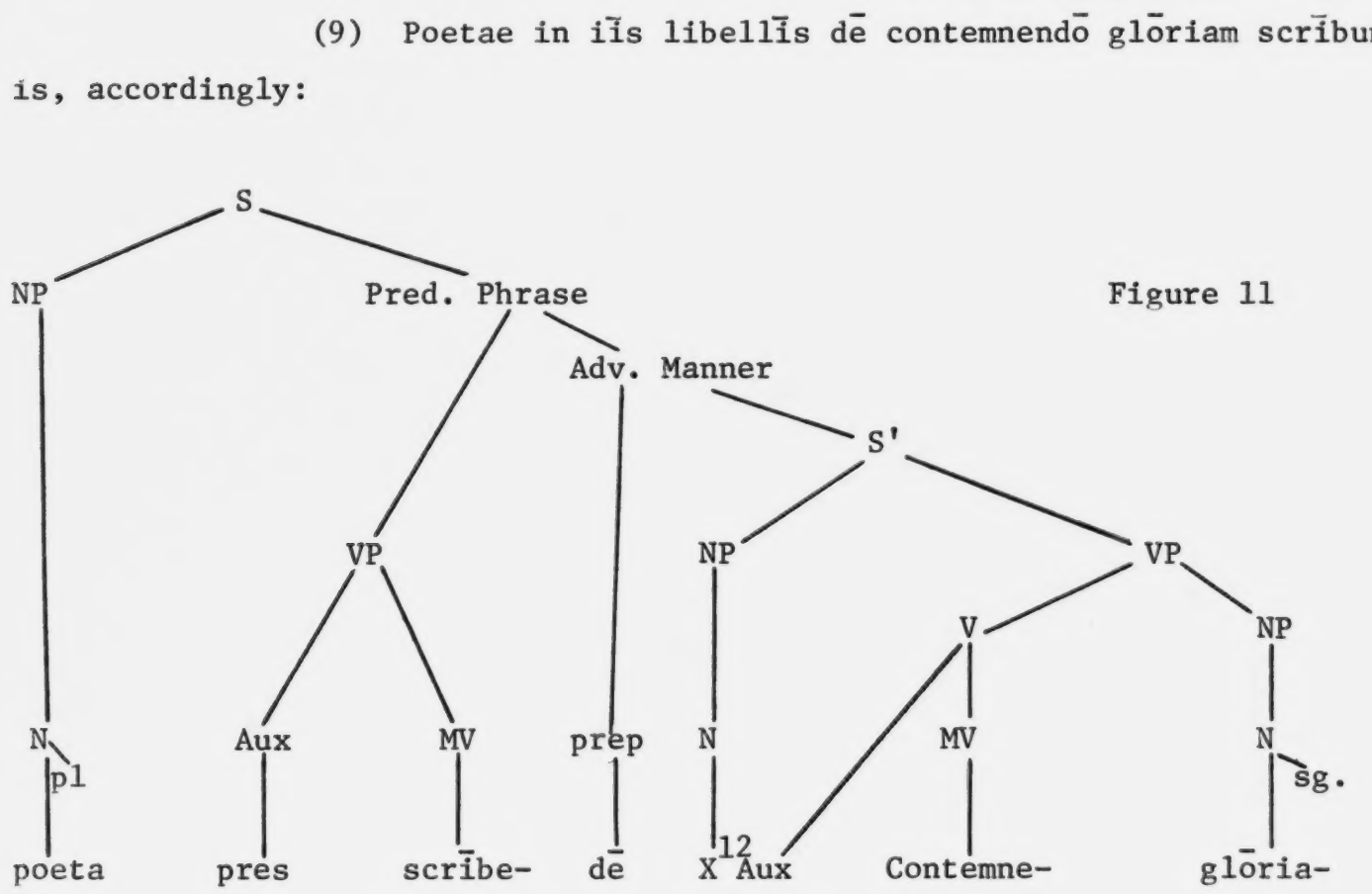


Figure 11

What Figure 11 specifies, in fact, is that the same embedded sentence belongs to the deep structure of both dē contemnendā gloria and dē contemnendō glōriam, that neither is passive, and, therefore, that gerundive constructions will not co-occur with agents. The overt shape of dē contemnendō glōriam is derived

transformationally by deletion of the subject NP which the embedded sentence shares with the matrix sentence, and replacement of Aux by ndo, which must be added to the Affix class, and will, therefore, be attached to the following verb stem by the Affix rule. The sequence V + ndo will automatically receive the case marker required by the preposition, in this case ablative. Dē contemnendā glōriā is, then, optionally transformed from dē contemnendō glōriam by the appropriate transformational rule. Notice that gerundives must be derived from gerunds, and not vice versa, since otherwise there will be gerund constructions which remain unaccounted for, i.e., gerunds whose underlying phrase markers contain intransitive verbs.

Many of the problems of the above examples, especially those in Latin, have been simplified. Although none of the basic features of their description (that would be necessary within the context of a complete syntax of Latin) have been changed. The purpose here was to show the necessity of distinguishing two different levels in the syntactical analysis of sentences: deep structure, which is not limited to the overt shape of sentences but which is more abstract and makes explicit the syntactical features and relations of each sentence; and superficial structure, which is derived from the former by transformational rules, mostly of deletion and re-arranging, and which is the basis of the phonological shape of sentences.

Footnotes

[1]The research reported herein was performed in part pursuant to Contract OEC-3-6-061784-0508 with the U. S. Department of Health, Education, and Welfare, Office of Education, under the provisions of P. L. 83-531, Cooperative Research, and the provisions of Title VI, P. L. 85-864, as amended. This research report is one of several which have been submitted to the Office of Education as Studies in Language and Language Behavior, Progress Report VIII, February 1, 1969.

[2]See, for example, the following: D. H. Kelly a) "The new linguistics and the teacher of classics," Classical Journal, 1966, 61, 247-251; b) "Modern linguistics and the teacher of Latin," Classical World, 1967, 60, 361-366; c) "Transformations in the Latin nominal phrase," Classical Philology, 1967, 63, 46-52; A. R. Keiler, Review of Twombly and O'Brien, A basic course in Latin, Language Learning, 1963, 13, 269-276.

[3]Most of these recent developments are discussed in N. Chomsky, Aspects of the theory of syntax. Cambridge: M.I.T. Press, 1965.

[4]A thorough analysis of pre-Aspects phrase structure rules can be found in P. Postal, Constituent structure: A study of contemporary models of syntactic description. The Hague: Mouton, 1964.

[5]The following description of English verbal expressions was first elaborated by Chomsky in Syntactic structures. The Hague: Mouton, 1957, 38-40.

[6]In order to facilitate the exposition I have chosen to focus attention on the differences in explicitness between phrase structure diagrams (or, following Aspects phrase markers which result from base rules) and phrase markers which result from transformational rules. It should be kept in mind throughout, however, that both sets of phrase markers result from different explicit types of rewriting rules, the purpose of which is just to assign such phrase markers to sentences.

[7]What is claimed here for Latin is simply that free word order cannot be the basis of the abstract level of representation that is part of the deep structure of sentences. That is, it is usually the case that some fixed order is necessary in order to be able to generalize in the appropriate way. It is still a difficult matter in many cases to know what that fixed underlying order has to be for 'free' word order languages. In the phrase markers for the remaining Latin examples, I have chosen what is generally assumed to be the neutral or unmarked word order, e.g., subject + object + verb rather than any of the other possibilities. I do not comment, however, any more about word order since it does not affect any of the grammatical arguments.

[8]I have omitted here, and in most of the following examples, some of the details which are required transformationally. These do not, however, pertain directly to the exposition.

[9]There is, of course, a personal passive construction available for verbs like dicō as well, e.g., mīlitēs dictī sunt castra pōnere, which is probably even more prevalent than the impersonal passive (although this is a matter of stylistic preference, for whatever reason, and not a matter of "grammatical preference"). There are, moreover, several ways in which this can be accounted

for: a) One might consider (4) to be structurally ambiguous, i.e., although analyzable into a single set of immediate constituents, susceptible to different semantic interpretations because of some syntactical facts not present in the surface form of the sentence. In other words, one could assume that it is like English examples such as Flying planes can be dangerous, where the choice of is/are in place of can will make the difference clear. This is, in effect, what Twombly and O'Brien do (cf. my review, op. cit). But this is clearly not the case. (4) is unambiguous quite apart from the fact that two passive constructions are available. b) It is, therefore, more plausible to derive the "personal"passive construction for sentences like (4) from the impersonal type. The latter type, of course, has the effect of bringing into relief the NP subject of the embedded sentence. Thus the availability of a "personal" passive construction for verbs of the dicō class is not counterevidence for the analysis given above.

[10]Kühner, R., & Stegmann, C. Ausführliche Grammatik der lateinischen Sprache. Hannover: 1966, 688.

[11]Halle, W. G., & Buck, C. D. A Latin grammar. New York: Mentzer, Bush & Co., 1903, p. 329.

[12]I mean by x here simply that the understood subject of the verb underlying the gerund (or gerundive) here is not the subject of the matrix sentence, i.e., poetae. The restriction on the subjects of matrix and embedded sentences underlying gerund and gerundive constructions is a problem that deserves close attention.